Evaluation of Defensive Linemen

2020 Sports Info Solutions Football Analytics Challenge

Nate Rowan

July 19, 2020

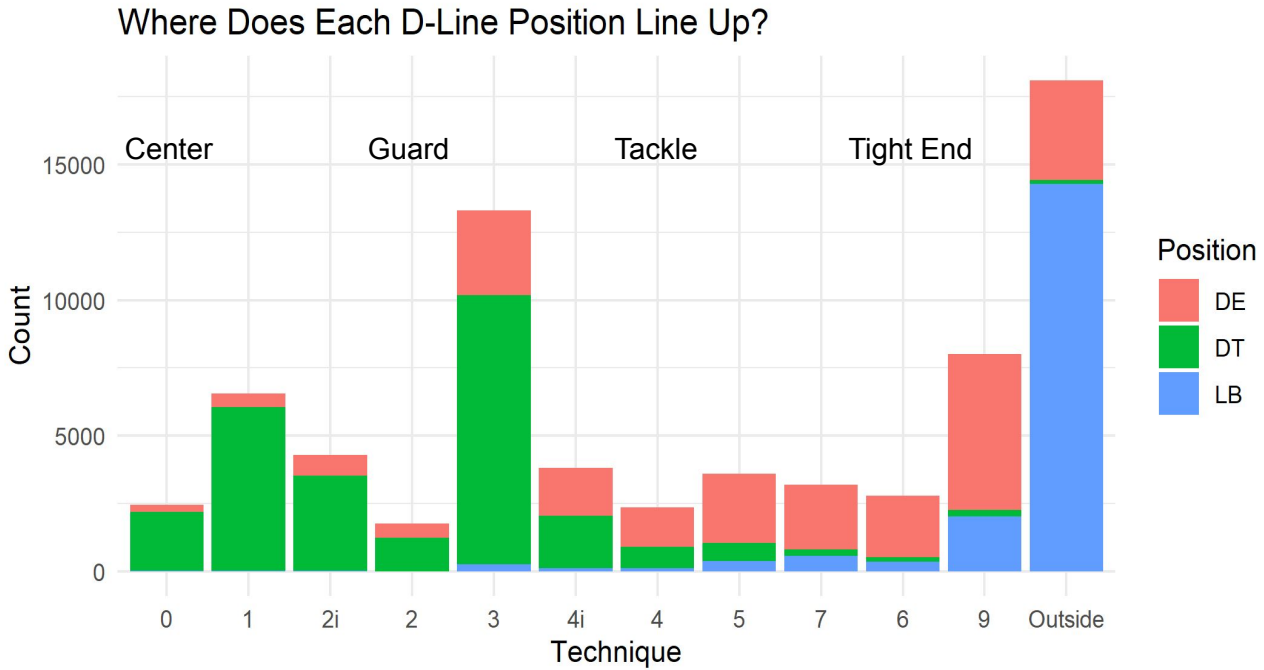
Outline

- Definitions
- Create a metric to evaluate player performance
- Apply that metric to each position
- Examine how findings change under different situations
- Limitations of the evaluation method

How will we define the defensive positions?

- I'll be using the provided RosterPosition variable, which indicates what position a player is listed at on the team's roster
- Positions: defensive end (DE), defensive tackle (DT), linebacker (LB)
    - I have filtered out players listed as defensive backs, and a couple of fullbacks who got defensive snaps

- Defensive tackles (green) almost always line up between the offensive tackles
- Linebackers (blue) almost always line up at least as wide as the tight end
- Defensive ends (red) are pretty spread out, but almost always outside the guards



What does our data look like?

- All data used in this project comes from weeks 9-17 of 2019 (no additional datasets were used)
- All QB kneels and spikes were removed for the analysis

How will we evaluate each player?

- Traditionally, this is done through counting stats: tackles, pressures, sacks, forced fumbles, etc
- Problem: what is the value of each statistic relative to one another?
  - A sack is more valuable than a pressure, but how much more valuable?
- Solution: calculate the difference between the average EPA of any play, and the average EPA of a play where a certain event happens
  - Similar to baseball's Linear Weights approach used to calculate Weighted On Base Average (wOBA)
    - Calculate the average increase in run expectancy created by a walk, single, double, etc
    - Multiply number of walks, singles, doubles, etc by the average increase in run expectancy of each outcome, and bam! A neat metric is created!
- We will look at this in two parts: pass plays and run plays
  - Similar to how Strokes Gained in golf evaluates driving, approach shots, short game, and putting separately

Pass Plays

Example: What is the value of a sack?

1. The average EPA for all passing plays in the data is -0.017
2. The average EPA for all sacks in the data is -1.415
3. So, a sack results in a difference in EPA of (-1.415 - -0.017) = -1.398. **A sack is worth about 1.4 points to the defense compared to the average passing play. We will refer to this as “Points Gained”**

We can repeat this process for pressures, forced fumbles, pass breakups, and interceptions to come up with a value for each statistic

Value of Each Counting Stat on Pass Plays

Stat	Points Gained	Observations
Pressure (no sack)	**0.098**	2868
Sack (no fumble)	**1.40**	567
Forced Fumble	**3.69**	81
Pass Breakup (no int)	**0.838**	973
Interception	**4.13**	203

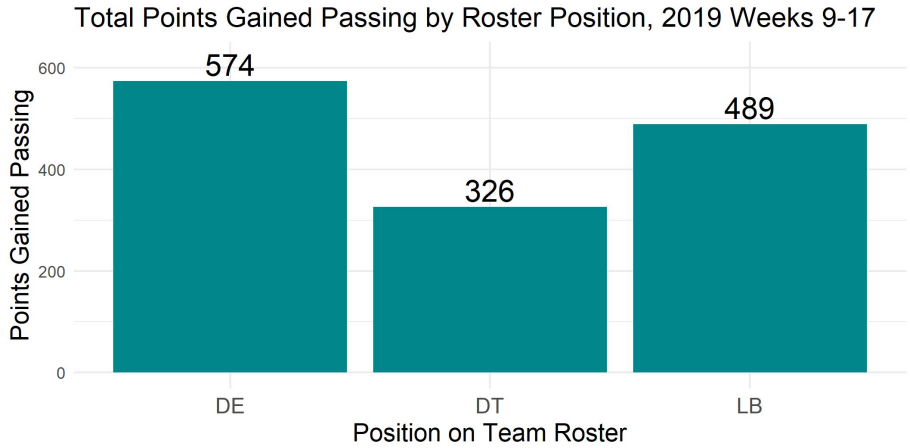
Points Gained Passing = .098*(pressures-sacks) + 1.40*(sacks-fumbles) + 3.76*ff + .838*(pbu-int) + 4.13*int

Why the subtractions? The “pressure” statistic includes sacks and forced fumbles, and the “sack” statistic includes forced fumbles. The 0.098 value for pressure is on pressures *where no sack/fumble occurs*

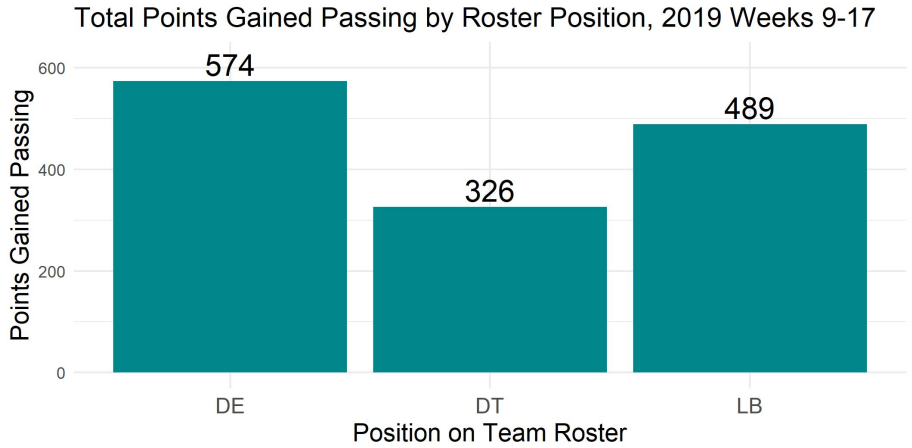
Now, we can easily calculate the Points Gained Passing of each player!

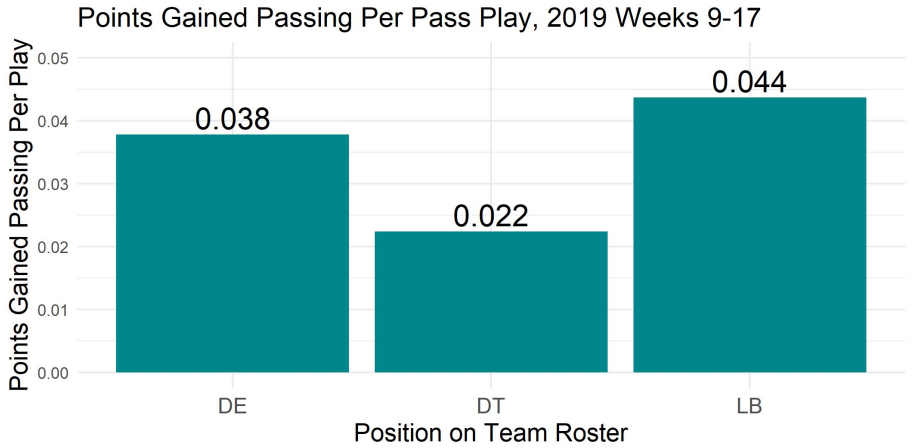
- Example: Chandler Jones
- In the provided dataset (second half of 2019), Jones had:
  - 37 pressures, 10 solo sacks, 1 assisted sack, 4 forced fumbles, 4 pass breakups, 0 interceptions
  - Now, need to subtract sacks from pressures and fumbles from sacks, giving us a stat line of: 26.5 pressures, 6.5 sacks, 4 forced fumbles, 4 pass breakups
- So, his Points Gained Passing for the second half of the season was: 0.098*26.5 + 1.40*6.5 + 3.76*4 + .838*4 + 4.14*0 = **29.78**
- **Chandler Jones gained about 30 points for the Cardinals on pass plays in the second half on 2019.**

Looking at Total Points Gained Passing, it appears that defensive ends were most valuable…



Looking at Total Points Gained Passing, it appears that defensive ends were most valuable...





...but when we account for the number of snaps at each position, linebackers overtake defensive ends

Run Plays

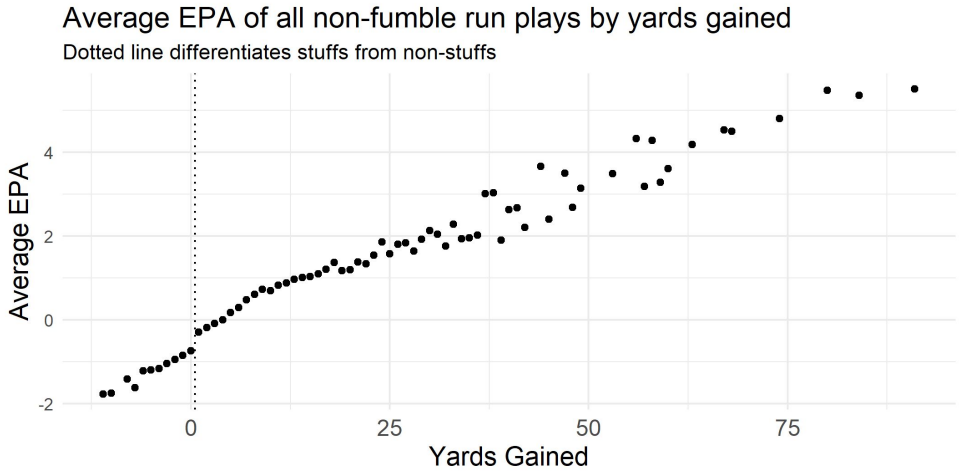
Much more difficult!

How can we quantify contributions to the run game?

- Total tackles are a poor way to measure defensive value
- We will look at two statistics: “stuffs” and forced fumbles
- Note: this method DEFINITELY undervalues contributions to the run game. I elaborate on this in the final slide.

Wait, what is a stuff, and why are we using them?

- I am defining a stuff as any run play that results in either 0 or negative yards
- They are something that we can easily measure, and are indicative of skill
- Additionally, using zero yards as a cutoff gives us a clear grouping (notice the jump in average EPA from 0 to 1 yard gained in the below plot)



What is the value of a stuff?

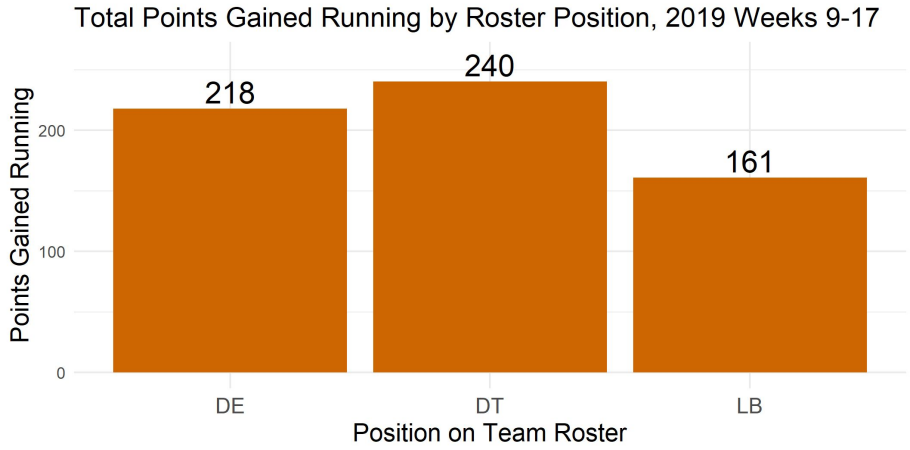
1. Average EPA of all run plays: -0.020
2. Average EPA of a stuff play: -0.858
3. So, a stuff results in a difference in EPA of -0.858 - (-0.020) = -0.838. **A stuff is worth about 0.8 points to the defense compared to an average run play**

Value of Each Rushing Counting Stat

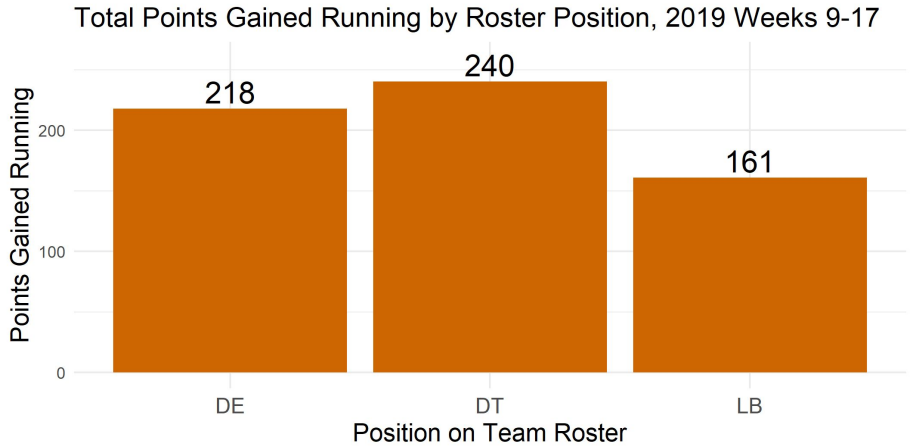
Stat	EPA Above Average	Observations
Stuff	**0.838**	1190
Forced Fumble	**2.371**	92

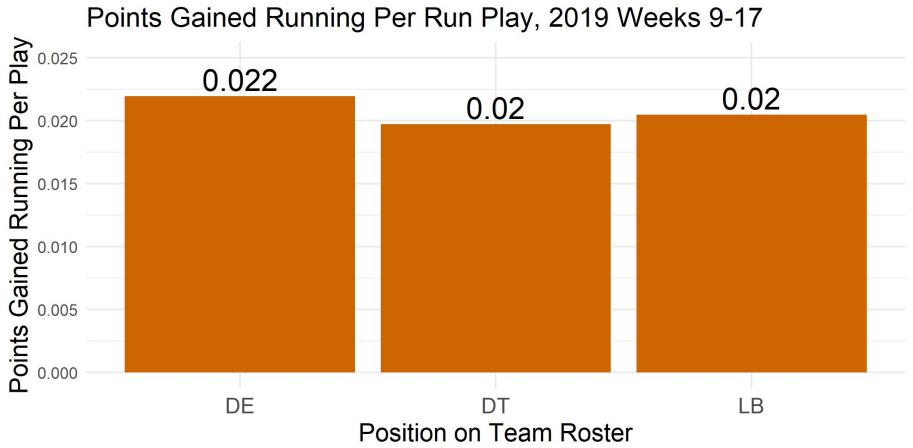
Points Gained Running = .838*stuffs + 2.371*fumbles

Looking at Total Points Gained Running, it appears that defensive tackles were most valuable…



Looking at Total Points Gained Running, it appears that defensive tackles were most valuable...



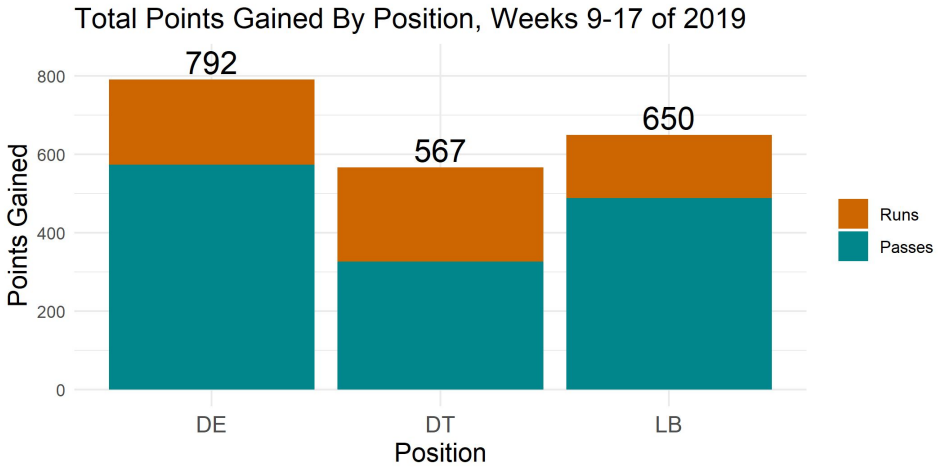


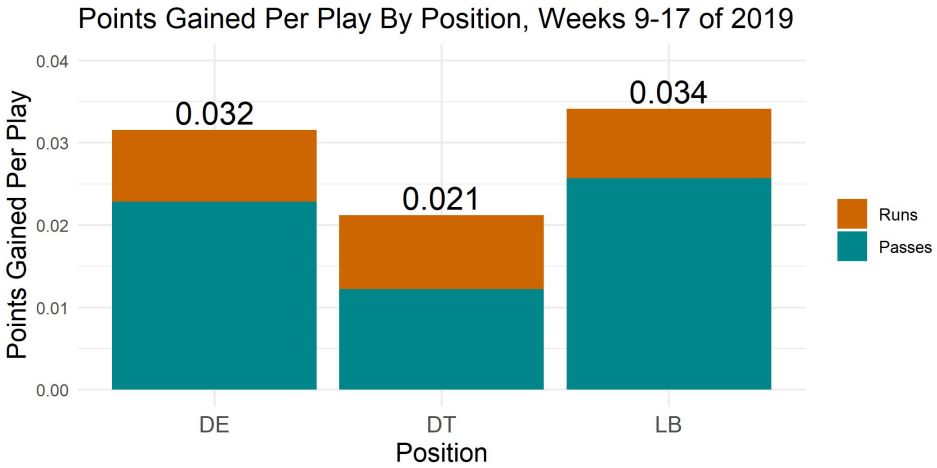
...but when we account for the number of snaps at each position, all positions are almost identical

Let’s Put it All Together!

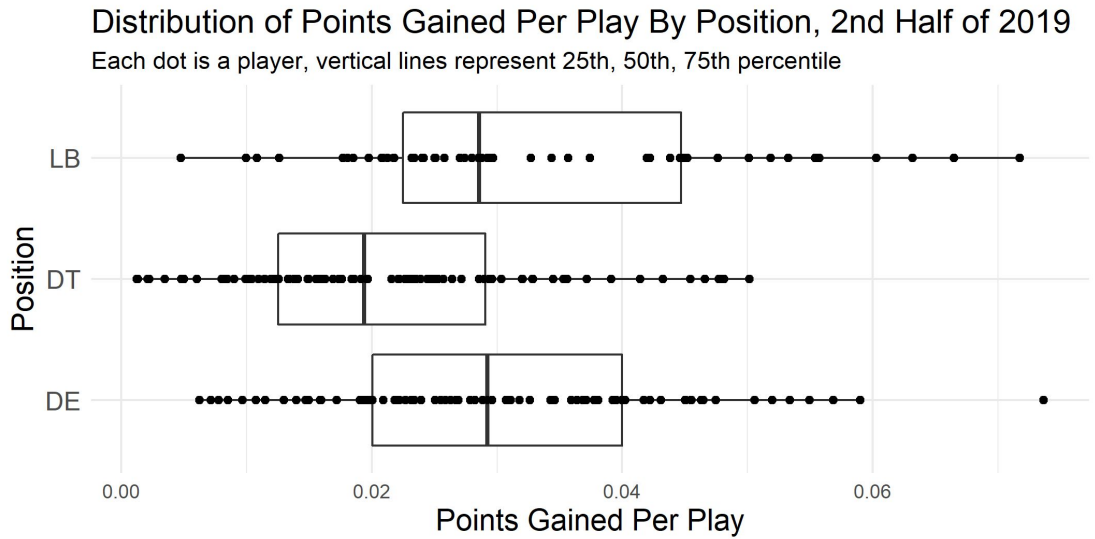
Total Points Gained

By raw total points gained, defensive ends are most valuable. If we look at a per snap basis, defensive ends and linebackers are about equal in value, with defensive tackles lagging behind





Distribution of Talent at Each Position



2 Main Takeaways:

- The median defensive end or linebacker is worth the same as the 75th percentile defensive tackle
- The difference between a top-tier LB/DE and a bottom-tier LB/DE is greater than the difference between a top-tier DT and a bottom-tier DT. So, an elite edge rusher provides more value over a bad edge rusher than an elite defensive tackle provides over a bad DT.

How does situation change our findings?

- On 3rd/4th down and short, “stuffs” on run plays become extremely valuable

Down and Distance	Average EPA, No Stuff	Average EPA, Stuff	Difference
2nd and 1	0.131	-0.758	0.89
3rd and 1	0.518	-1.57	2.09
4th and 1	1.70	-3.26	4.96

- On 4th and 1, a stuff results in nearly a 5 point swing! Teams ran the ball about 75% of the time on 4th and 1 in the dataset, so an effective run-stopper becomes much more valuable than an effective edge rusher in this situation

Summary of Conclusions

- Overall, linebackers and defensive ends appear to be more valuable on a per-snap basis than defensive tackles
    - In the run game, however, all positions seem to be equally valuable
- The difference between an elite edge rusher and a replacement-level edge rusher is greater than the difference between an elite defensive tackle and a replacement-level defensive tackle.
- On 3rd/4th and short, the value of run-stoppers dramatically increases

Limitations/Flaws of this analysis

1. Run plays are definitely undervalued, harming defensive tackles
    a. Players can contribute in ways outside of stuffs and forced fumbles. A tackle for a one yard gain is valuable. Filling your gap is valuable, even if it doesn't result in a tackle.
    b. This article from Sharp Football Analysis provides a way to evaluate defensive players on runs that seems more thorough.
2. "Plays that don't show up in the stat sheet" aren't valued by the Points Gained metric
    a. For example, an elite pass rusher who is double-teamed every play might not accumulate a ton of sacks, but still provides value.
3. The Points Gained Per Play rate stat is flawed
    a. The denominator I used is total snaps.
    b. However, this probably isn't a great denominator. Just because a player is on the field doesn't mean they have an opportunity to make a play. In baseball, plate appearances are an awesome denominator, because each PA is an opportunity to make a play. It's extremely difficult to define opportunities for defensive linemen

Thank you!

I had a blast working on this project, thank you so much for setting up the challenge, providing the data, and taking the time to read through this presentation!